

Observação: em todos os problemas que envolvem teste de hipótese, é necessário deixar claro e responder nesta ordem:

- As hipóteses nula (H_0) e alternativa (H_a)
- O valor crítico
- A estatística de teste
- A decisão do teste (rejeita ou não rejeita H_0)
- Uma frase escrevendo a conclusão com base na decisão

1. Os diâmetros de uma população de rolamentos fabricados por uma empresa tem distribuição normal com média de 1 mm e desvio padrão de 0,2 mm. Qual é a probabilidade de:

- (a) Amostrar 100 rolamentos e obtermos um diâmetro médio maior que 1,13 mm?

```
## score z
(ez <- z(xbarra = 1.13, mu = 1, sigma = 0.2, n = 100))
[1] 6.5
## probabilidade
round(pnorm(ez, lower.tail = FALSE), 3)
[1] 0
```

Ou $0,0001^1$

- (b) Amostrar 5 rolamentos e obtermos um diâmetro médio menor que 0,93 mm?

```
## score z
(ez <- z(xbarra = 0.93, mu = 1, sigma = 0.2, n = 5))
[1] -0.78
## probabilidade
round(pnorm(ez, lower.tail = TRUE), 3)
[1] 0.218
```

- (c) Amostrar 35 rolamentos e obtermos um diâmetro médio superior a 1,15 mm ou inferior a 0,97 mm?

```
## score z
(ez1 <- z(xbarra = 1.15, mu = 1, sigma = 0.2, n = 35))
[1] 4.44
(ez2 <- z(xbarra = 0.97, mu = 1, sigma = 0.2, n = 35))
[1] -0.89
## probabilidade
round(pnorm(ez1, lower.tail = FALSE) + pnorm(ez2, lower.tail = TRUE), 3)
[1] 0.187
```

- (d) Refaça o item (b), mas agora com uma amostra de 50 rolamentos. Compare as respostas e explique a diferença.

```
## score z
(ez <- z(xbarra = 0.93, mu = 1, sigma = 0.2, n = 50))
[1] -2.47
## probabilidade
round(pnorm(ez, lower.tail = TRUE), 3)
[1] 0.007
```

A probabilidade de obter um diâmetro médio menor do que 0,93 mm é bem maior quando temos uma amostra pequena (caso do item b). Quando aumentamos o tamanho da amostra (caso do item d), a probabilidade de encontrarmos um diâmetro médio menor que 0,93 mm diminui. Isso ocorre pois conforme aumentamos o tamanho da amostra, aumentamos também a quantidade de informação referente ao fenômeno estudado, aproximando cada vez mais a estimativa do verdadeiro valor da população.

2. A quantidade de horas diárias que um adulto passa em frente ao computador possui uma distribuição com assimetria positiva, com média de 3,6 horas e desvio padrão de 0,5 horas. Qual é a probabilidade de:

- (a) Avaliarmos 87 pessoas e obtermos uma média superior a 3,7 horas?

¹NOTA: Neste exercício, o valor crítico de $z_{\alpha/2}$ foi de 6,5, um valor maior do que é mostrado na tabela. Para qualquer valor acima de 3,99 (último valor dado na tabela), podemos usar uma área de 0,49999, pois sabemos que a área total à direita (ou esquerda) da média 0 vale 0,5, no entanto, os valores de $z_{\alpha/2}$ variam entre $-\infty$ e ∞ .

```
## escore z
(ez <- z(xbarra = 3.7, mu = 3.6, sigma = 0.5, n = 87))
[1] 1.87
## probabilidade
round(pnorm(ez, lower.tail = FALSE), 3)
[1] 0.031
```

(b) Avaliarmos 6 pessoas e obtermos uma média inferior a 2,9 horas?

```
## escore z
(ez <- z(xbarra = 2.9, mu = 3.6, sigma = 0.5, n = 6))
[1] -3.43
## probabilidade
round(pnorm(ez, lower.tail = TRUE), 4)
[1] 3e-04
```

(c) Avaliarmos 37 pessoas e obtermos uma média superior a 3,67 horas ou inferior a 3,47 horas?

```
## escore z
(ez1 <- z(xbarra = 3.67, mu = 3.6, sigma = 0.5, n = 37))
[1] 0.85
(ez2 <- z(xbarra = 3.47, mu = 3.6, sigma = 0.5, n = 37))
[1] -1.58
## probabilidade
round(pnorm(ez1, lower.tail = FALSE) + pnorm(ez2, lower.tail = TRUE), 3)
[1] 0.255
```

(d) O que pode-se dizer com relação ao resultado obtido no item (b)?

Como a amostra é pequena e a distribuição dos dados é assimétrica, não podemos usar o resultado do item (b) para se fazer inferência.

3. Se queremos construir intervalos de confiança para as estimativas de médias populacionais, quais seriam os valores de α e críticos de $z_{\alpha/2}$ se estamos interessados nos níveis de confiança de:

- (a) 90% (b) 94% (c) 95% (d) 98% (e) 99%

```
##-----
## (a) 90%
qnorm(0.1/2, lower.tail = FALSE)
[1] 1.644854
##-----
## (b) 94%
qnorm(0.06/2, lower.tail = FALSE)
[1] 1.880794
##-----
## (c) 95%
qnorm(0.05/2, lower.tail = FALSE)
[1] 1.959964
##-----
## (d) 98%
qnorm(0.02/2, lower.tail = FALSE)
[1] 2.326348
##-----
## (e) 99%
qnorm(0.01/2, lower.tail = FALSE)
[1] 2.575829
```

4. Retiramos uma amostra aleatória de 55 preços de um determinado produto. A média amostral foi de R\$ 18,30, e assume-se que o desvio padrão populacional é conhecido, com valor de R\$ 4,10. Com isso:

(a) Verifique se as suposições necessárias para o cálculo de intervalos de confiança estão satisfeitas.

AAS ✓ σ conhecido ✓ $n > 30$ ✓

(b) Calcule as margens de erro para os níveis de confiança

(i) $\gamma = 0,90$ (ii) $\gamma = 0,95$ (iii) $\gamma = 0,99$

```
##-----
## (i) 0,90
```

```
icz(alfa = 0.1, xbarra = 18.3, sigma = 4.1, n = 55)
```

```
$`valor critico`
```

```
[1] 1.644854
```

```
$`margem de erro`
```

```
[1] 0.9093473
```

```
##-----
```

```
## (i) 0,95
```

```
icz(alfa = 0.05, xbarra = 18.3, sigma = 4.1, n = 55)
```

```
$`valor critico`
```

```
[1] 1.959964
```

```
$`margem de erro`
```

```
[1] 1.083554
```

```
##-----
```

```
## (i) 0,99
```

```
icz(alfa = 0.01, xbarra = 18.3, sigma = 4.1, n = 55)
```

```
$`valor critico`
```

```
[1] 2.575829
```

```
$`margem de erro`
```

```
[1] 1.424031
```

- (c) Construa os intervalos de confiança para os três casos do item anterior.

```
##-----
```

```
## (i) 0,90
```

```
icz(alfa = 0.1, xbarra = 18.3, sigma = 4.1, n = 55, ic = 1)
```

```
$`intervalo de confiança`
```

```
[1] 17.39065 19.20935
```

```
##-----
```

```
## (i) 0,95
```

```
icz(alfa = 0.05, xbarra = 18.3, sigma = 4.1, n = 55, ic = 1)
```

```
$`intervalo de confiança`
```

```
[1] 17.21645 19.38355
```

```
##-----
```

```
## (i) 0,99
```

```
icz(alfa = 0.01, xbarra = 18.3, sigma = 4.1, n = 55, ic = 1)
```

```
$`intervalo de confiança`
```

```
[1] 16.87597 19.72403
```

- (d) Explique o que significa cada intervalo de confiança obtido.

(Exemplo de resposta para 90% de confiança) Temos 90% de confiança de que o intervalo entre 17,39 e 19,21 contém o verdadeiro preço médio deste produto.

-
5. Por analogia a produtos similares, o tempo de reação de um novo medicamento pode ser considerado como tendo distribuição Normal com desvio-padrão igual a 2 minutos (a média é desconhecida). Vinte pacientes foram sorteados, receberam o medicamento e tiveram seu tempo de reação anotado. A média do tempo de reação dos 20 pacientes amostrados foi de 4,75 minutos.

- (a) Construa um intervalo de confiança de 90% para estimar a verdadeira média para o tempo de reação.

```
icz(alfa = 0.1, xbarra = 4.75, sigma = 2, n = 20, ic = 2)
```

```
$`valor critico`
```

```
[1] 1.644854
```

```
$`margem de erro`
```

```
[1] 0.7356009
```

```
$`intervalo de confiança`
```

```
[1] 4.014399 5.485601
```

- (b) Construa um intervalo de confiança de 99% para estimar a verdadeira média para o tempo de reação.

```
icz(alfa = 0.01, xbarra = 4.75, sigma = 2, n = 20, ic = 2)
```

```
$`valor critico`
```

```
[1] 2.575829
```

```
$`margem de erro`  
[1] 1.151946
```

```
$`intervalo de confiança`  
[1] 3.598054 5.901946
```

- (c) Compare e interprete os resultados. O que você pode afirmar quanto à influência do aumento do nível de confiança na amplitude dos intervalos?

Amplitude do intervalo de 90%: $AMP_{90} = 5,486 - 4,014 = 1.472$ e amplitude do intervalo de 99%: $AMP_{99} = 5,902 - 3,598 = 2.304$. O intervalo de 90% de confiança possui uma amplitude menor do que o intervalo de 99% de confiança. Isso ocorre pois conforme aumentamos a confiança de que um intervalo de valores realmente posso conter o parâmetro populacional desconhecido, naturalmente o intervalo de valores deve ser maior. Note que a margem de erro aumenta quando aumentamos a confiança, o que significa que quanto mais confiança desejamos ter em um intervalo, maior será o erro que estamos assumindo.

6. Seja $X \sim N(\mu, 16)$

- (a) Uma amostra de tamanho 25 foi coletada e forneceu uma média amostral de 8. Construa intervalos de confiança de 90% e 95%.

```
##-----  
## gama = 0,9  
icz(alfa = 0.1, xbarra = 8, sigma = 4, n = 25, ic = 2)  
$`valor critico`  
[1] 1.644854
```

```
$`margem de erro`  
[1] 1.315883
```

```
$`intervalo de confiança`  
[1] 6.684117 9.315883
```

```
##-----  
## gama = 0,95  
icz(alfa = 0.05, xbarra = 8, sigma = 4, n = 25, ic = 2)  
$`valor critico`  
[1] 1.959964
```

```
$`margem de erro`  
[1] 1.567971
```

```
$`intervalo de confiança`  
[1] 6.432029 9.567971
```

- (b) Para um nível de confiança de 95%, construa intervalos de confiança supondo tamanhos de amostra de 15 e 65 (admita a mesma média amostral de 8). Comente sobre o porque os intervalos de confiança são diferentes.

```
##-----  
## n = 15  
icz(alfa = 0.05, xbarra = 8, sigma = 4, n = 15, ic = 2)  
$`valor critico`  
[1] 1.959964
```

```
$`margem de erro`  
[1] 2.024242
```

```
$`intervalo de confiança`  
[1] 5.975758 10.024242
```

```
##-----  
## n = 65  
icz(alfa = 0.05, xbarra = 8, sigma = 4, n = 65, ic = 2)  
$`valor critico`  
[1] 1.959964
```

```
$`margem de erro`  
[1] 0.9724145
```

```
$`intervalo de confiança`  
[1] 7.027586 8.972414
```

- (c) Calcule o tamanho da amostra, para que com 90% de probabilidade, a média amostral não se afaste da média populacional por mais de (i) 0,2 unidades, (ii) 2 unidades. Comente sobre o porque os tamanhos amostrais são diferentes.

```
##-----  
## (i) e = 0,2  
nam(0.1, sigma = 4, erro = 0.2)  
[1] 1083  
##-----  
## (ii) e = 2  
nam(0.1, sigma = 4, erro = 2)  
[1] 11
```

-
7. Desejamos coletar uma amostra de uma variável aleatória X , com distribuição Normal, de média desconhecida e variância 30. Qual deve ser o tamanho da amostra para que, com 95% de probabilidade, a média amostral não difira da média populacional por mais de 3 unidades?

```
nam(0.05, sigma = sqrt(30), erro = 3)  
[1] 13
```

-
8. Considere que a altura de árvores em uma floresta é o objeto de interesse. Temos interesse em fazer inferências sobre a verdadeira altura média populacional das árvores. Alguns monitoramentos são feitos para a coleta de dados sobre as alturas das árvores. Após uma análise descritiva verificamos que a distribuição das alturas das árvores segue uma distribuição normal. Se temos interesse em estimar intervalos de confiança para a média verdadeira das alturas:
- (a) Qual seria a distribuição de probabilidade apropriada para a análise da distribuição das médias amostrais?
 - (b) Quais as suposições necessárias para construirmos intervalos de confiança quando o σ é desconhecido?
 - (c) Quais seriam os valores de α e críticos se estivéssemos interessados em construir intervalos de confiança de 99%, 95% e 90% para a média verdadeira com uma amostra de 38 alturas de árvores.
 - (d) Quais seriam os valores críticos apropriados se estivéssemos interessados em construir intervalos de confiança de 99%, 95% e 90% para a média verdadeira com uma amostra de 13 alturas de árvores.

-
9. No caso do problema anterior, estime as margens de erro das médias amostrais em relação à verdadeira média se:
- (a) O tamanho amostral é 14 e queremos margens de erro associadas a 95% de confiança. Na amostra encontramos uma média de 1,3 m e um desvio padrão de 0,4 m.
 - (b) O tamanho amostral é 65 e queremos margens de erro associadas a 95% de confiança. Na amostra encontramos uma média de 2 m e um desvio padrão de 1 m.
 - (c) O tamanho amostral é 28 e queremos margens de erro associadas a 99% de confiança. Na amostra encontramos uma média de 0,6 m e um desvio padrão de 0,2 m.

-
10. Quais são os intervalos de confiança para a verdadeira média populacional para os itens (a), (b) e (c) da questão anterior? Escreva uma conclusão à respeito de cada intervalo de confiança obtido.

-
11. Uma pesquisa de consumo de combustível realizada entre 61 clientes de um posto indicou um consumo médio de 11,2 km/l, com desvio-padrão de 2,8 km/l. Suponha distribuição normal para o consumo médio. Com isso:
- (a) Determine um intervalo de confiança de 95% para o consumo médio populacional dos clientes.

```
ict(alfa = 0.05, xbarra = 11.2, s = 2.8, n = 61, ic = 2)  
$`valor critico`  
[1] 2.000298
```

```
$`margem de erro`  
[1] 0.7171133
```

```
$`intervalo de confiança`  
[1] 10.48289 11.91711
```

- (b) Assuma que foram avaliados apenas 9 clientes. Calcule intervalos de confiança para a média ($\bar{x} = 11,2$) para os seguintes casos
 - i. Assumindo σ desconhecido ($s = 2,8$)

```
ict(alfa = 0.05, xbarra = 11.2, s = 2.8, n = 9, ic = 2)
```

```
$`valor critico`
```

```
[1] 2.306004
```

```
$`margem de erro`
```

```
[1] 2.152271
```

```
$`intervalo de confianca`
```

```
[1] 9.047729 13.352271
```

ii. Assumindo σ conhecido ($\sigma = 2,8$)

```
icz(alfa = 0.05, xbarra = 11.2, sigma = 2.8, n = 9, ic = 2)
```

```
$`valor critico`
```

```
[1] 1.959964
```

```
$`margem de erro`
```

```
[1] 1.8293
```

```
$`intervalo de confianca`
```

```
[1] 9.3707 13.0293
```

Qual a sua conclusão referente à diferença observada entre os dois intervalos de confiança? Por que ocorre essa diferença?

Notamos que quando o desvio-padrão populacional (σ) é desconhecido, o intervalo de confiança obtido com o uso da distribuição t é maior (em amplitude), do que o intervalo obtido através da distribuição normal padrão Z . Esta diferença ocorre porque o valor crítico obtido através da distribuição t é maior do que o valor crítico da distribuição Z . O valor crítico da distribuição t leva em conta o tamanho da amostra, que neste caso é pequeno ($n < 30$), além do fato de não se conhecer o σ verdadeiro. Com isso, a distribuição t tende a gerar intervalos de confiança maiores para amostras pequenas.

12. Você deseja estimar a quantia média de taxas anuais pagas por estudantes de universidades privadas. É razoável supor que estas quantidades variam entre R\$ 0,00 (por exemplo, para estudantes com bolsa integral) até R\$ 40.000,00. Determine o número de estudantes que devem ser selecionados para termos 95% de confiança de que a média amostral esteja no máximo a:

(a) R\$ 100,00 da verdadeira média populacional

```
nam(alfa = 0.05, sigma = (40000 - 0)/4, erro = 100)
```

```
[1] 38415
```

(b) R\$ 500,00 da verdadeira média populacional

```
nam(alfa = 0.05, sigma = (40000 - 0)/4, erro = 500)
```

```
[1] 1537
```

O que você pode dizer à respeito das diferenças entre os tamanhos amostrais obtidos?

A medida que aumentamos o erro amostral, ou seja, quando assumimos que a nossa estimativa da média amostral \bar{x} esteja mais afastada da verdadeira média populacional μ (desconhecida), então o tamanho da amostra é menor. Quando diminuimos a margem de erro da estimativa que estamos interessados (ou seja, assumimos errar menos), então necessariamente precisamos coletar uma amostra maior, para conseguirmos mais informação sobre o objeto em estudo, e assim garantir uma maior precisão da estimativa.

13. O que é a probabilidade de cometer um erro do tipo I, e o que significa esse termo?

O erro do tipo I é a probabilidade de rejeitar a hipótese nula, quando ela é de fato verdadeira, ou seja,

$$P(\text{rejeitar } H_0 | H_0 \text{ verdadeira}) = \alpha$$

Esta probabilidade é representada por α , e deve ser determinada antes do início do teste de hipótese. Este valor de α , ou o nível de significância do teste determina o quanto “aceitamos errar” quando realizamos um teste de hipótese.

14. Uma máquina é projetada para fazer esferas de aço de 1 cm de raio. Uma amostra de 10 esferas é produzida, e tem raio médio de 1,004 cm, com $s = 0,003$. Há razões para suspeitar que a máquina esteja produzindo esferas com raio maior que 1 cm? (Use um nível de significância de 10%).

```
$`valor critico`  
[1] 1.383029
```

```
$`estatistica de teste`  
[1] 4.21637
```

```
$decisao  
[1] "rejeita H0"
```

15. O desvio-padrão da vida útil de um monitor de LED é $\sigma = 500$, e a distribuição da vida útil é considerada normal. O fabricante afirma que a vida útil média é, no mínimo 9000 horas. Testar esta afirmação ao nível de significância de 5%, dado que em uma amostra de 15 televisores de LED, a média foi de 8800 horas.

```
$`valor critico`  
[1] -1.644854
```

```
$`estatistica de teste`  
[1] -1.549193
```

```
$decisao  
[1] "nao rejeita H0"
```

16. Uma fábrica de automóveis anuncia que seus carros consomem, em média, 11 litros por 100 km, com desvio-padrão de 0,8 litro. Uma revista decide testar essa afirmação e analisa 35 carros dessa marca, obtendo 11,4 litros por 100 km como consumo médio. Admitindo que o consumo tenha distribuição normal, ao nível de 10%, o que a revista concluirá sobre o anúncio da fábrica?

```
$`valor critico`  
[1] -1.644854 1.644854
```

```
$`estatistica de teste`  
[1] 2.95804
```

```
$decisao  
[1] "rejeita H0"
```

17. Um investidor interessado em desenvolver um centro comercial, é informado por um representante de um grupo comunitário que a renda média familiar na comunidade é de no mínimo R\$ 15.000,00 por ano. Supõe-se que os valores de renda na população sejam normalmente distribuídos. Para uma amostra aleatória de $n = 15$ famílias, a média amostral é de R\$ 14.000,00, e o desvio-padrão amostral é de R\$ 2.000,00. Testar a hipótese do salário médio na comunidade ser no mínimo R\$ 15.000,00, ao nível de 5% de significância.

```
$`valor critico`  
[1] -1.76131
```

```
$`estatistica de teste`  
[1] -1.936492
```

```
$decisao  
[1] "rejeita H0"
```

18. Cinco medidas do conteúdo de alcatrão em um cigarro acusaram $\bar{x} = 14,4$ mg e $s = 0,274$ mg. Pretende-se testar a hipótese de que a média de alcatrão nos cigarros é igual a 14,1 mg, conforme declarado na embalagem, ao nível de 5% de significância. Qual seria a conclusão do teste quando as hipóteses alternativas forem:

(a) $H_a \neq 14,1$ (b) $H_b < 14,1$ (c) $H_c > 14,1$

(a)

```
$`valor critico`  
[1] -2.776445 2.776445
```

```
$`estatistica de teste`  
[1] 2.44825
```

```
$decisao
```

```
[1] "nao rejeita H0"
```

```
#### (b)
```

```
$`valor critico`
```

```
[1] -2.131847
```

```
$`estatistica de teste`
```

```
[1] 2.44825
```

```
$decisao
```

```
[1] "rejeita H0"
```

```
#### (c)
```

```
$`valor critico`
```

```
[1] 2.131847
```

```
$`estatistica de teste`
```

```
[1] 2.44825
```

```
$decisao
```

```
[1] "rejeita H0"
```

-
19. Um pesquisador deseja estudar o efeito de certa substância no tempo de reação de seres vivos a um certo tipo de estímulo. Um experimento é desenvolvido com cobaias que são inoculadas com a substância e submetidas a um estímulo elétrico. Os tempos de reação, em segundos, foram: 9,1; 9,3; 7,2; 7,5; 13,3; 10,9; 7,2; 9,9; 8,0; 8,6. Admite-se que o tempo de reação segue, em geral, o modelo normal com média 8 e desvio-padrão $\sigma = 2$. O pesquisador desconfia, entretanto, que o tempo médio sofre alteração por influência da substância. Faça um teste de hipótese para, ao nível de 6% de significância, para verificar a desconfiança do pesquisador.

```
$`valor critico`
```

```
[1] -1.880794 1.880794
```

```
$`estatistica de teste`
```

```
[1] 1.739253
```

```
$decisao
```

```
[1] "nao rejeita H0"
```